

Effects of Food Assistance and Nutrition Programs on Nutrition and Health: Volume II, Data Sources

Christopher Logan, Mary Kay Fox, and Biing-Hwan Lin

Introduction

The Nutrition and Health Outcomes Study is being carried out by Abt Associates Inc., under contract with the Economic Research Service (ERS) of the U.S. Department of Agriculture (USDA). One goal of the study was to assess the potential for using available and future data sources in evaluating the impact of USDA's food assistance and nutrition programs (FANPs) on nutrition- and health-related outcomes. The FANPs of interest are the Food Stamp Program (FSP), the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), the National School Lunch Program (NSLP), and the School Breakfast Program (SBP). The information presented in this report is an inventory of data sources that can be used as a foundation for future FANP-related research.

This report presents assessments of data sources of potential interest to researchers in ERS and elsewhere and identifies those most worthy of future investigation. The report focuses exclusively on documenting what is available in extant data sources or expected to be available through ongoing data collection programs. No attempt has been made to define or recommend, in more than a general way, analyses that could be completed with the available data. Moreover, it should be noted that, in describing available data, we assumed that researchers may be interested in using these data sources for exploratory or descriptive analyses as well as impact analyses. Consequently, we cast a very wide net in identifying potentially interesting variables. For example, we documented the availability of measures of cognitive and emotional status and performance.

A caveat to using any of the available data sources to measure program impacts is that such analyses are likely to be subject to selection bias, although the analysis may be specifically structured to address this problem. As emphasized in the first report prepared under this contract, statistical adjustments may be used to deal with the problem of selection bias, but one can never be certain that bias has been successfully removed (Hamilton and Rossi, 2001).

The report is divided into three sections. The rest of this introductory section provides a brief overview of the approach used to identify the data sources we examined, as well as a description of the methods used to obtain and compile the information for assessing the usefulness of each source. Subsequent sections describe and compare the principal and potential data sources.

Identifying Data Sources for Review

Identification and review of data sources were conducted mainly in mid-1999, but information known to be in flux was updated periodically until this report was finalized. The search assumed that the ideal source would combine the following characteristics:

- coverage of both program participants and eligible nonparticipants for at least one of the programs in FANPs;¹
- identification of program participation status and sufficient information to determine eligibility for nonparticipants; and

¹ Some approaches to evaluating program impacts do not require data on nonparticipants (e.g., dose-response models), but the most useful sources have such data. See Hamilton and Rossi (2001) for further discussion.

- nutrition- or health-related measures that might be useful in studying program impacts or in describing characteristics of program participants and nonparticipants.

To compile a comprehensive inventory of available data, we cast a wide net in identifying potentially useful variables for FANP research.

The search considered available databases that might meet all of these criteria, as well as ongoing or planned data collection programs that could do so in the future. Development of the list of potential data sources began with the data sources used in the National Nutrition Monitoring and Related Research Program (NNMRRP) (Federation of American Societies for Experimental Biology (FASEB), 1995). Additional sources were identified through a careful review of major database references, as well as through searches of World Wide Web sites of government agencies and other research organizations known to collect nutrition- and/or health-related data. To ensure relevance to the current policy and program environment, databases that included only data collected before 1990 were not considered for review.

Recently, the Interagency Board for Nutrition Monitoring and Related Research (IBNMRR, 2000) compiled a list of surveys, studies, and databases that contain data on (1) nutrition- and health-related measurements, (2) food and nutrient consumption, (3) knowledge, attitudes, and behavior assessments, (4) food composition and nutrition databases, (5) food supply determinations, and (6) nutrition monitoring activities in the States. This document was not reviewed in detail, but a preliminary check (as well as the authors' and reviewers' knowledge) suggested that no important sources have been overlooked. Many of the surveys listed in the IBNMRR report are not reviewed here because they do not meet the aforementioned criteria, but they may contain data useful for research pertaining to nutrition, diet, and health.

Completing Initial Utility Assessments

A standard profile form was used to document each data source. The profile included the following information:

- purpose of research,
- sponsoring organization/agency,

- data collection timeline (beginning and end dates),
- population covered,
- sampling design,
- FANPs for which participation is identified,
- nutrition- and health-related data,
- demographic data,
- data availability, and
- purchasing/ordering contact and costs.

For the most promising sources, FANP participation variables and nutrition and health measures were documented in more detail.

Information on the characteristics of each data source was obtained from the first volume of the Third Report on Nutrition Monitoring in the United States (FASEB, 1995); searches of World Wide Web sites; and telephone and mail communication with sponsoring agencies and organizations. Copies of survey questionnaires and associated documentation were obtained from the World Wide Web, Abt Associates' in-house resources, or the data repository for the source.

Categorizing the Relevance of Data Sources

As noted, it was assumed that the ideal data source for FANP research would include coverage of both program participants and eligible nonparticipants for at least one FANP; identification of program participation status and sufficient information to determine eligibility for nonparticipants; and potentially useful nutrition- and/or health-related measures.

Using these criteria, each data source was assigned to one of four categories reflecting current and potential usefulness for FANP research, described below.

Principal Sources

This category includes existing databases that appear to have the greatest potential for conducting FANP research. They meet all three of the criteria noted above and, in some cases, provide the best or only source of data for a given outcome. All but one of the principal data sources are recurrent data collection programs, offering opportunity for historical analyses, and several are longitudinal in design. The principal sources have undergone the most detailed of the assessments for this report.

Potential Sources

The data sources in this category are less useful than the principal sources because they cover limited populations, have a critical gap in participation information (for example, they include birth outcome measures but not information on WIC participation), or include outcome measures that, by comparison, are weaker than those available in the principal sources. (Some potential sources have limitations in two or more of these areas.) Many of these sources are ongoing data collection programs that could be made more useful for future FANP research if questions were incorporated to fill gaps in relevant information.

Recognized Sources

This category includes existing databases that have already been used to evaluate one or more of the FANPs. These sources are either dated (including only pre-1990 data) or have not been expanded since their initial analyses were conducted, so it is unlikely that additional analyses are warranted. Recognized sources are listed in table 1.

Insufficient Sources

Data sources assigned to this category were not considered candidates for future analyses and detailed assessments were not completed. Although many of these sources include relevant data, they were judged to have very little potential for evaluating outcomes of FANPs. Sources were classified as insufficient for having one or more of the following drawbacks:

- including only pre-1990 data;
- having limited or nonexistent measures of interest (program participation and/or outcomes), with the gaps either impossible to fill (extant databases) or unlikely to be rectified (narrowly focused ongoing data collection program);
- including a very limited population (descriptive studies of program participants are in this category);
- being proprietary; or
- being an administrative database that would require a complex linkage to another data source to combine information on program participation and outcomes.

A list of insufficient data sources is given in Appendix A.

It should be noted that the classification of administrative databases as insufficient is based on the purposes of this study. Nevertheless, these data sources can be merged with other databases that provide information on program participation. A good example is the WIC-Medicaid database constructed for USDA's Food and Nutrition Service (FNS) in the late 1980s (Devaney et al., 1991).

The next sections provide detailed descriptions of principal and potential data sources, respectively. A detailed profile of each source is provided in Appendix B.

Principal Data Sources

The data sources described in this section are currently available datasets that can be used to examine the effects of FANPs on nutrition and health outcomes. Each source meets the minimum requirements described in the preceding section. In addition, all but two of the principal data sources are ongoing national data collection programs that are updated on a regular schedule, allowing update of any analyses as subsequent data become available.² The recurring nature of these data collection programs also provides the opportunity for future expansion of program participation information and/or nutrition and health outcome measures. Some data collection programs have been in progress for some time, offering the potential for longitudinal analyses at the aggregate or individual level. Finally, some of the principal data sources represent the best or only source of information for certain outcomes of interest.

Table 2 summarizes key characteristics of the principal data sources. Sources are listed in alphabetical order, by title. The table shows the major FANPs for which participation information is available and the broad categories of nutrition- and health-related measures that are available.³ It also provides the time period for

² As discussed later in this section, two of the principal sources have been combined into a single ongoing program since the most recent available data were collected.

³ Table 2 focuses on participation data for the Special Supplemental Nutrition Program for Women, Infants and Children Program (WIC), the Food Stamp Program (FSP), the National School Lunch Program (NSLP), and the School Breakfast Program (SBP). These are the four FANPs most often identified in major surveys. Available participation data for other FANPs (quite limited by comparison) are discussed in the text.